

CLAIMS

What is claimed is:

1. A method of matching sequences of signals from an information stream comprising:
 - 5 extracting a stream of samples from a subject information stream, each of the extracted samples comprising a portion of the information stream;
 - computing, for each of the samples, a vector quantity indicative of the data in the sample;
 - correlating the vectors to generate a signature indicative of the stream of samples;
 - 10 comparing the generated signatures to signatures generated from a match stream of samples; and
 - generating, as a result of the comparing, a distance matrix indicative of signatures generated from similar samples.
- 15 2. The method of claim 1 wherein correlating includes generating the signature distinctively from signatures generated for other streams of samples.
3. The method of claim 1 wherein the correlating generates a covariance matrix as a signature indicative of the stream of samples.
4. The method of claim 1 further comprising arranging the samples into a plurality of segments.
- 20 5. The method of claim 1 wherein each of the extracted samples comprises a predetermined interval of the subject information stream.

6. The method of claim 1 wherein the match stream is the extracted stream of samples itself.
7. The method of claim 1 wherein the match stream is indicative of a predetermined sequence of samples stored from previously transmitted samples.
- 5 8. The method of claim 7 wherein signatures corresponding to the predetermined sequence are stored in a library.
9. A method of detecting repetitions in an information stream comprising:
 - providing an information stream of multimedia data corresponding to a transmitted signal;
 - 10 extracting a plurality of samples from the information stream;
 - accumulating the samples into segments comprising a predetermined interval of the transmitted signal;
 - generating respective vectors indicative of the samples in each of the segments;
 - 15 for each segment, correlating the vectors in the segments to generate a respective covariance matrix corresponding to the segment;
 - aggregating each of the covariance matrices corresponding to the segments into a sequence of covariance matrices; and
 - 20 comparing each of the covariance matrices in the sequence with each other covariance matrix in the sequence to generate a distance matrix.
10. The method of claim 9 wherein generating the distance matrix further comprises generating a distance value, indicative of the similarity between the distance matrices, as a result of the comparing.

11. The method of claim 9 further comprising traversing the distance matrix to determine similar sequences of covariance matrices.
12. The method of claim 11 wherein determining similar sequences further comprises searching for diagonals of similar distance values.
- 5 13. The method of claim 12 wherein searching for diagonals of similar distance values includes searching for distance values within a predetermined threshold.
14. The method of claim 9 wherein the covariance matrices correspond to a multidimensional space and comparing the covariance matrices further comprises:
10 determining, for each of the covariance matrices, a location in the multidimensional space; and
 computing a distance value indicative of the distance between the covariance matrices in the multidimensional space.
- 15 15. The method of claim 14 further comprising ranking the distance values and storing the ranking in the distance matrix.
16. The method of claim 15 wherein a matching sequence is determined by similar diagonal values in the distance matrix.
17. The method of claim 16 wherein the similar diagonal values form a contiguous sequence for a predetermined number of locations in the distance matrix.
- 20 18. The method of claim 14 wherein the multidimensional space further comprises at least 39 dimensions.

19. The method of claim 9 wherein the information stream further comprises streaming audio or video.
20. The method of claim 9 further comprising:
- 5 traversing the distance matrix to determine if a similar distance value sequence is found;
- querying, when a similar distance value sequence is found in the distance matrix, a library of previously found distance value sequences;
- storing the distance value sequence in the library if a match is not found as a result of the querying; and
- 10 updating, if a match is found as a result of the querying, a timestamp corresponding to the matching distance value sequence in the library.
21. The method of claim 20 further comprising
- traversing the distance value sequences in the library;
- examining each of the timestamps of the distance value sequences; and
- 15 purging, when the timestamp is beyond a predetermined threshold; the distance value sequence corresponding to the timestamp.
22. A system for detecting repetitions in an information stream comprising:
- 20 a stream processor operable to extract a stream of samples from the subject information stream and further operable to arrange the samples into a plurality of segments, each of the samples comprising a portion of the information stream;
- a segment processor operable to receive the segments and to compute, for each sample in the segments, a vector indicative of the data in the sample;
- 25 a correlator operable to correlate each of the vectors in at least one of the segments to generate a signature indicative of the segment; and

a distance processor responsive to the correlator and operable to compare the signature to signatures generated from a match stream of samples, and further operable to generate, as a result of the comparing, a distance matrix indicative of signatures generated from similar samples in the match stream.

- 5 23. The system of claim 22 wherein the signature generated by the correlator is distinctive from signatures generated from other samples.
24. The system of claim 23 wherein the signature is a covariance matrix.
25. The system of claim 22 wherein each of the segments comprises a predetermined interval of the information stream.
- 10 26. The system of claim 22 wherein the match stream is the stream of samples itself.
27. The system of claim 22 wherein the match stream is indicative of a predetermined sequence of samples stored from previously transmitted samples.
28. The system of claim 27 wherein the signature corresponding to the predetermined sequence is stored in a library.
- 15 29. A computer program product having computer program code for matching sequences of signals from an information stream comprising:
 - computer program code for extracting a stream of samples from a subject information stream, each of the extracted samples comprising a portion of the information stream;
 - 20 computer program code for computing, for each of the samples, a vector quantity indicative of the data in the sample;

computer program code for correlating the vectors to generate a signature indicative of the stream of samples;

computer program code for comparing the generated signatures to signatures generated from a match stream of samples; and

5 computer program code for generating, as a result of the comparing, a distance matrix indicative of signatures generated from similar samples.

30. A computer data signal including program code for matching sequences of signals from an information stream comprising:

10 program code for extracting a stream of samples from a subject information stream, each of the extracted samples comprising a portion of the information stream;

program code for computing, for each of the samples, a vector quantity indicative of the data in the sample;

15 program code for correlating the vectors to generate a signature indicative of the stream of samples;

program code for comparing the generated signatures to signatures generated from a match stream of samples; and

20 program code for generating, as a result of the comparing, a distance matrix indicative of signatures generated from similar samples.

31. A system for detecting repetitions in an information stream comprising:

means for extracting a stream of samples from a subject information stream, each of the extracted samples comprising a portion of the information stream;

25 means for computing, for each of the samples, a vector quantity indicative of the data in the sample;

means for correlating the vectors to generate a signature indicative of the stream of samples;

means for comparing the generated signatures to signatures generated from a match stream of samples; and

means for generating, as a result of the comparing, a distance matrix indicative of signatures generated from similar samples.